National Aeronautics and Space Administration

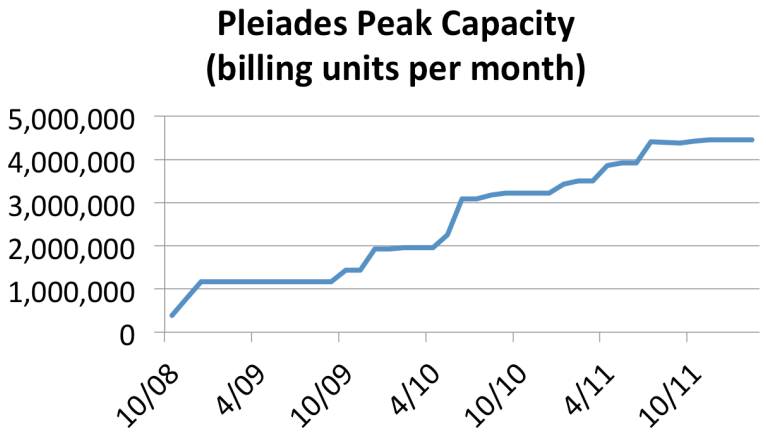
Tracking Pleiades Node Performance from Harpertown to Sandy Bridge

What is an SBU and why should I care?

www.nasa.gov

Pleiades Installation History

- 2008: 5888 nodes of Harpertown,
  512 nodes of Clovertown
  - Clovertowns isolated as separate machine soon after
- 2010: added 1280 nodes of Nehalem
- 2010-11: added 4672 nodes of Westmere
  - 64 nodes of Harpertown removed
- 2012: added 1728 nodes of Sandy Bridge
  - 1728 nodes of Harpertown removed



How to get users to send jobs to "best" node type for their app?

Question? Use the Webex chat facility to ask the Host

Standard Billing Unit (SBU)

- Charge for resource usage to reflect "computational power" of node types being used
  - e.g. faster, more recent machines have higher charging rates
- Establishing the charging rates:
  1. Define a representative suite of apps to reflect overall system workload
     - Including dataset, number of ranks, iteration count, etc.
  2. Run each of the apps on each of the node types
  3. Determine how many runs of each app can be made on a fixed amount of each node type—fractions OK
  4. For each node type, determine relative number of runs made compared to the *baseline* (Westmere node)
  5. For each node type, compute the weighted average across all apps of the relative number of runs
     - This is the *SBU charging rate* for that node type
     - Note: baseline node has rate of 1.0 (i.e. 1 SBU = 1 Westmere node-hour)

Question? Use the Webex chat facility to ask the Host

SBU Rate Example

- Two applications: A_1, A_2
- Two node types: N_1, N_2 (say N_1 is baseline)
- Run times:

	# nodes	N_1	N_2
A_1	64	1000	1500
A_2	128	500	1000

- Relative number of runs:

	N_1	N_2
A_1	1.0	0.67
A_2	1.0	0.50

- Weighted average:

	weight	N_1	N_2
A_1	30%	0.3	0.20
A_2	70%	0.7	0.35
SBU rate		**1.0**	**0.55**

Question? Use the Webex chat facility to ask the Host

Suite to Represent Pleiades' Workload

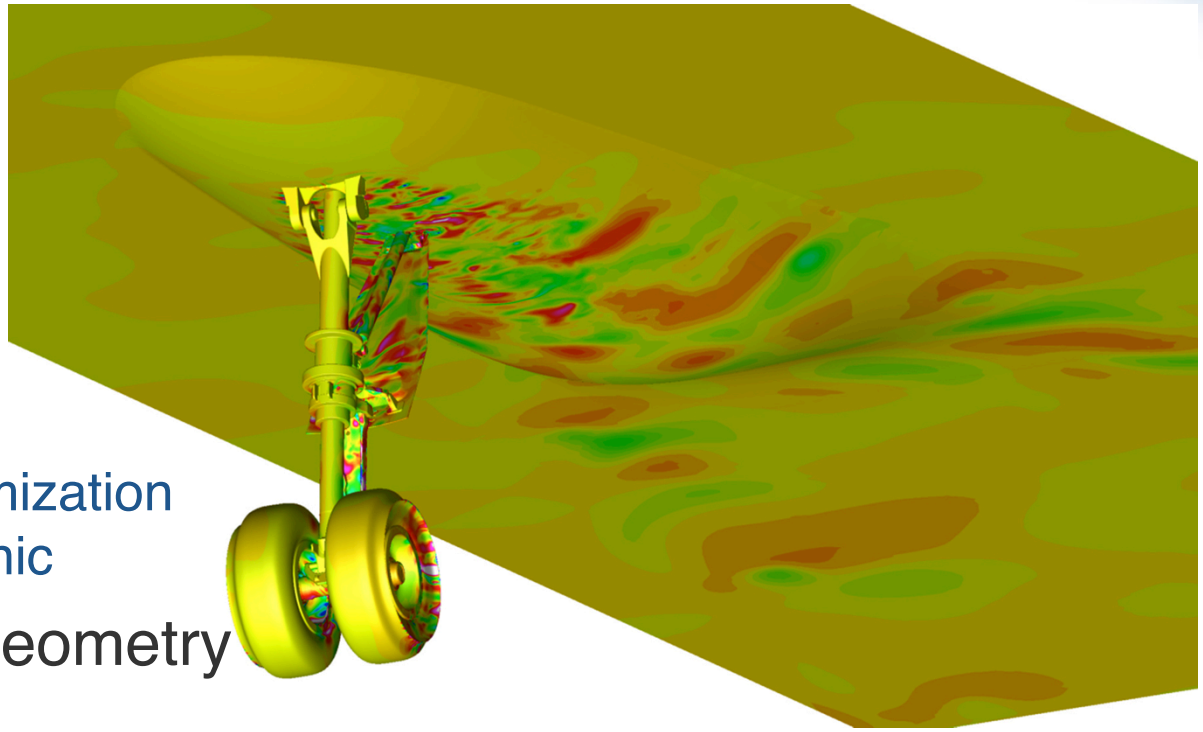
Application	# cores	Weight	NASA Mission Directorates
FUN3D	960	10%	Aeronautics; Human Exploration
OVERFLOW	480	20%	Aeronautics; Human Exploration
USM3D	480	20%	Aeronautics; Human Exploration
Enzo	240	20%	Science (non-Earth Science)
GEOS-5	1176	15%	Earth Science
WRF	384	15%	Earth Science

- Codes and weights chosen to represent the workload on Pleiades
- Core counts determined by finding "sweet spot" of scaling behavior
  - also, divisible by both 8 and 12
- Iteration counts adjusted so runs are ~1800 seconds using baseline (Westmere) nodes

Question? Use the Webex chat facility to ask the Host

FUN3D (v11.3)

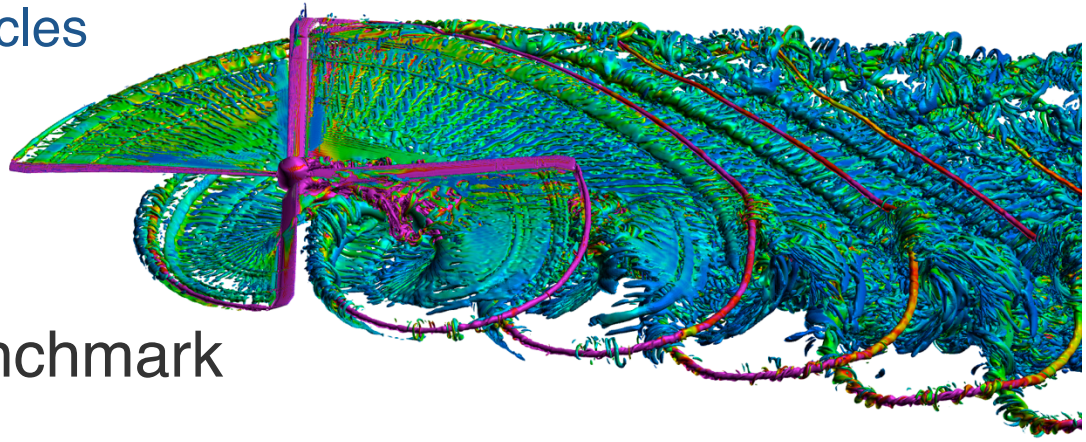
- Unstructured CFD code from NASA Langley
  - Adjoint-based error estimation
  - Mesh adaptation
  - Aerospace design optimization extending into hypersonic
- Dataset: wing-body geometry with transonic wing
  - Developed as a Common Research Model for validation studies of CFD codes
  - 3-D unstructured tetrahedra with 100M nodes



Question? Use the Webex chat facility to ask the Host

OVERFLOW (v2.1ae)

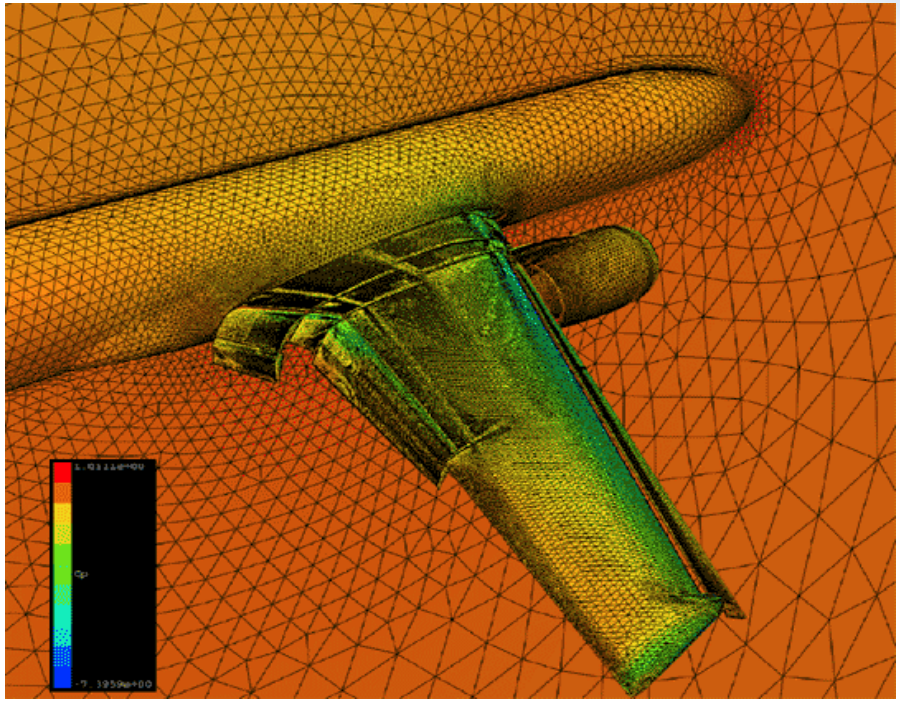
- Overset grid CFD program from NASA Langley
  - Launch and re-entry vehicles
  - Rotorcraft
  - Ships
  - Commercial aircraft
- Dataset: “nasrotor” benchmark geometry
  - 3-blade generic rotor system similar to UH-60 system
  - Constant NACA0010 airfoil section, rectangular planform
  - ~99M grid points



Question? Use the Webex chat facility to ask the Host

USM3D (v20100611)

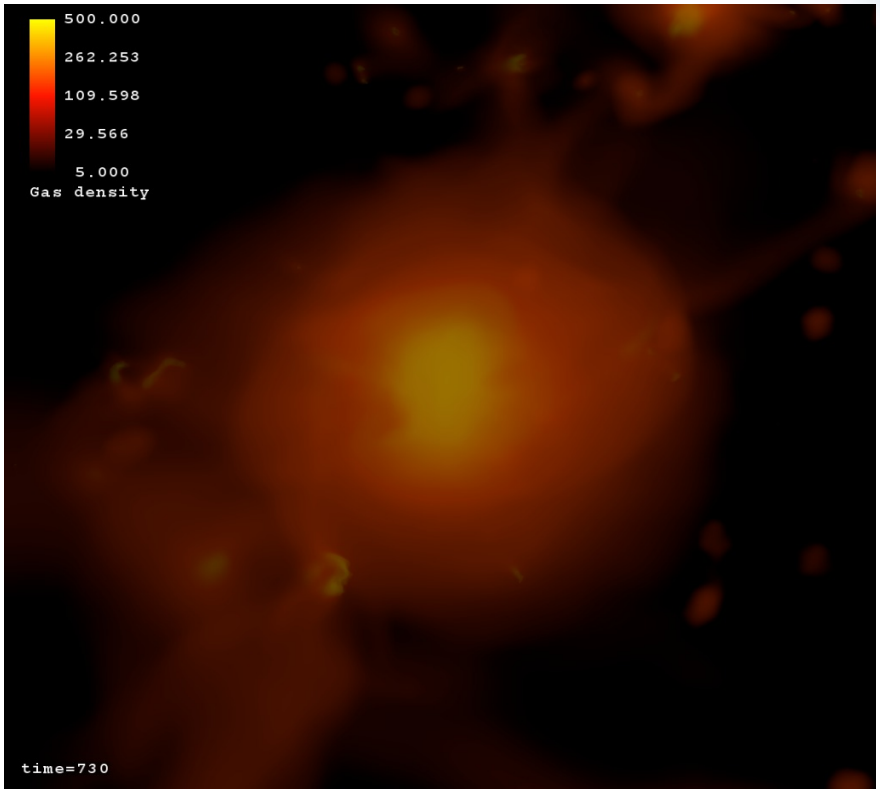
- Unstructured mesh CFD code from NASA Langley
  - Calculate flows over complex geometries
  - Aerodynamic flow of aerospace vehicle design
- Dataset: wing-body geometry with transonic wing
  - (same as in FUN3D)
  - Geometry used in a workshop in August 2008 to compare several state-of-the-art CFD codes



Question? Use the Webex chat facility to ask the Host

Enzo (v2.0)

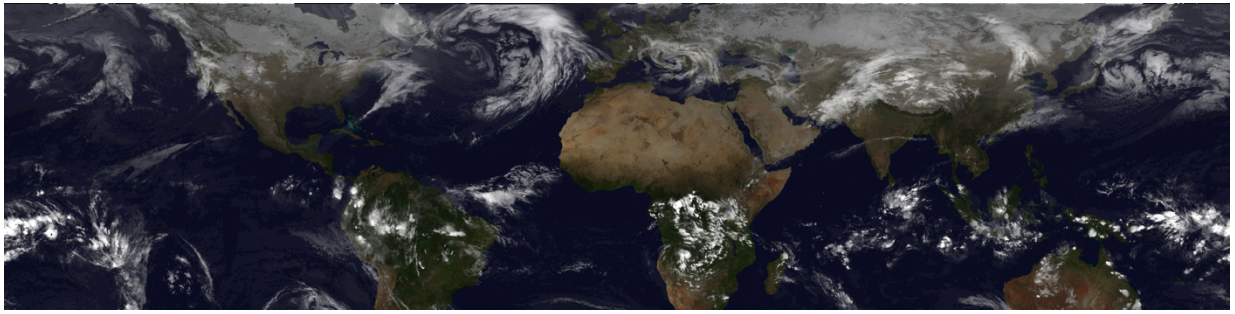
- Used to simulate cosmological structure formation
  - Grid-based with adaptive mesh refinement
- Dataset:
  - HDF5-format files represent initial cosmological conditions
  - Simulation then evolves from the initial conditions



Question? Use the Webex chat facility to ask the Host

GEOS-5

- Goddard Earth Observing System Model (v5)
  - Atmospheric general circulation model from NASA Goddard
  - Basic research
  - Data analysis
  - Climate and weather prediction
  - Observing system model and design

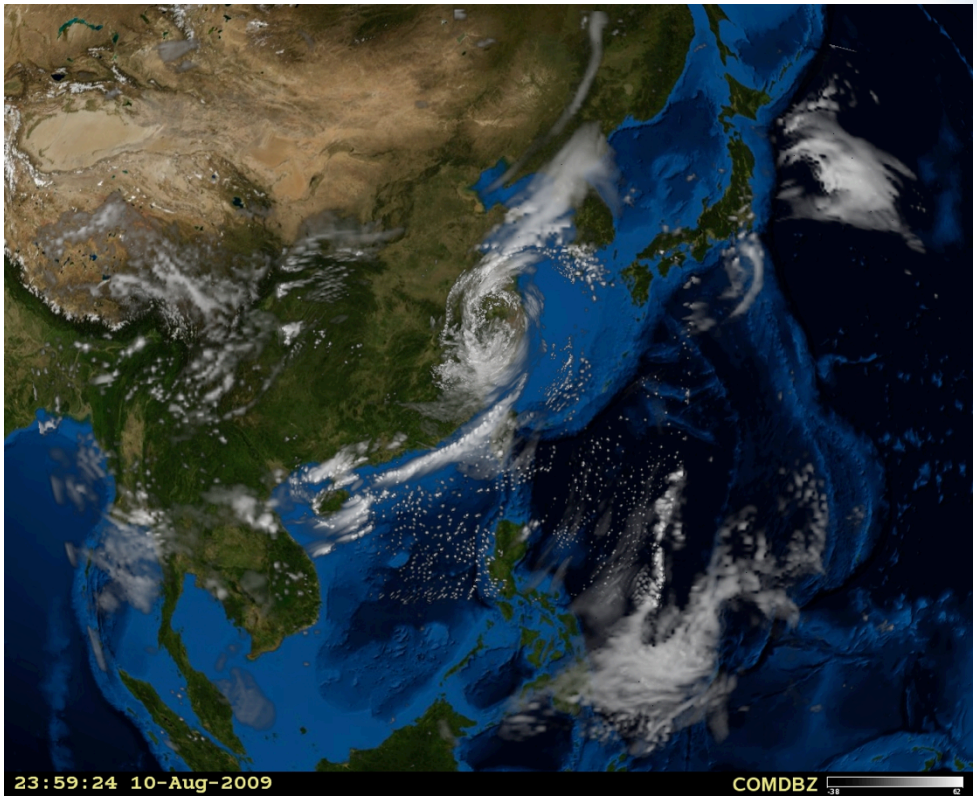


- Dataset: benchmark case 4 of the GEOS-5 model
  - Initial data generated by code itself
  - Resolution of 7km
  - Physical problem solves Jablonowski & Williamson Baroclinic Test Case

Question? Use the Webex chat facility to ask the Host

WRF (v3.1)

- Weather Research and Forcasting Model
  - Next-generation mesoscale numerical weather prediction from National Center for Atmospheric Research
  - Operational forecasting
  - Atmospheric research
- Dataset: Typhoon Morakot (SE Asia), August 2009
  - Generated from a previous calculation of WRF
  - Resolution of 2km



Question? Use the Webex chat facility to ask the Host

Timings

Application	Harpertown	Nehalem	Westmere	Sandy Bridge*	
				w/ TB*	w/o TB
FUN3D	3270	1616	1734	1405	1502
OVERFLOW	2931	1364	1786	1154	1253
USM3D	3954	1596	1802	1499	1670
Enzo	2128	1525	1925	1637	1839
GEOS-5	2815	1636	2096	1219	1361
WRF	3404	1775	2036	1499	1663

*TB=Turbo Boost

- Keep in mind: Westmere nodes have 12 cores instead of 8
  - Thus Westmere runs use only 2/3 of the nodes of a Harpertown or Nehalem run
  - The Harpertown or Nehalem time must be less than 2/3 of the Westmere time to be more efficient in resource utilization (none are)
- Similarly, the Sandy Bridge nodes have 16 cores

To compare node efficiencies, use the *relative number of runs* metric

Question? Use the Webex chat facility to ask the Host

Relative Run Numbers

Application	Harpertown	Nehalem	Westmere	Sandy Bridge	
FUN3D	0.35	0.71	1.0	1.65	1.54
OVERFLOW	0.41	0.87	1.0	2.06	1.90
USM3D	0.30	0.75	1.0	1.60	1.44
Enzo	0.60	0.68	1.0	1.57	1.40
GEOS-5	0.50	0.85	1.0	2.29	2.05
WRF	0.40	0.76	1.0	1.81	1.63

Weighed averages of the run numbers lead to these SBU rates:

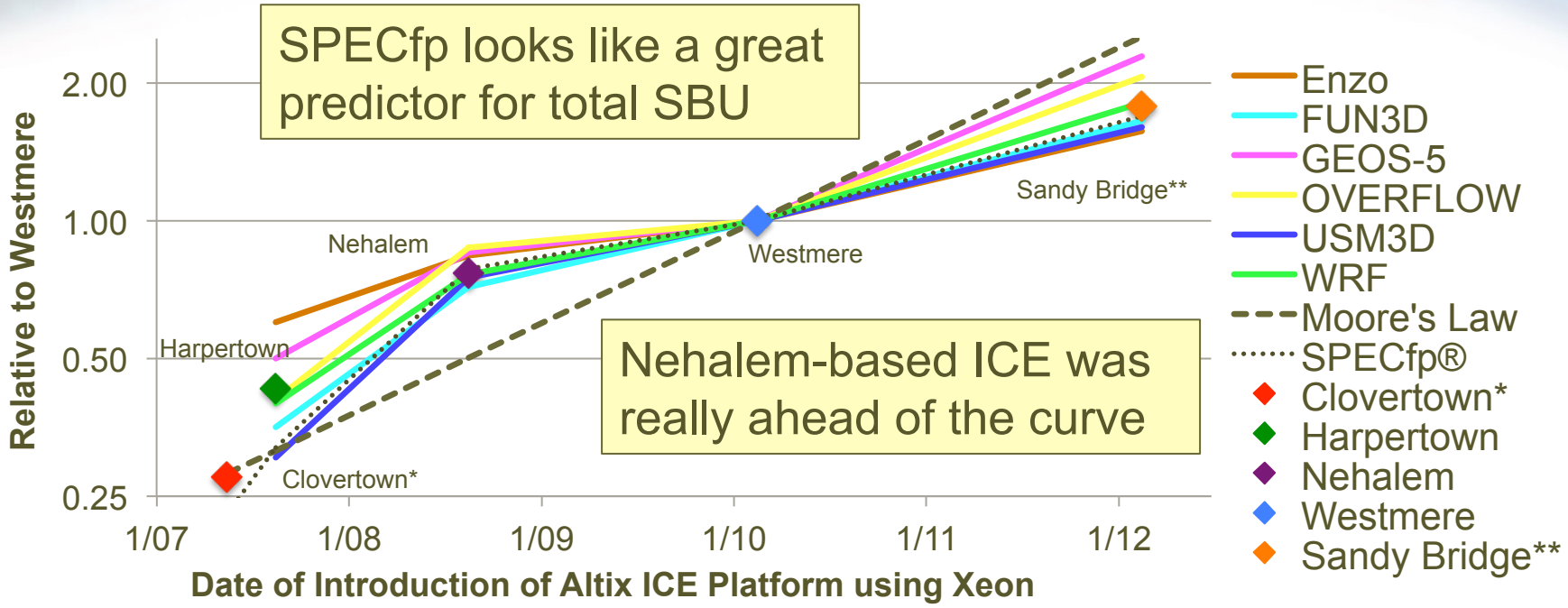
Application	Harpertown	Nehalem	Westmere	Sandy Bridge	
SBU rate	0.43	0.77	1.0	1.83	1.65

- Enzo does very well on Harpertown; GEOS-5 does not
- OVERFLOW does well on Nehalem; Enzo does not
- GEOS-5, OVERFLOW love Sandy Bridge; Enzo, USM3D not so much

(So far) WRF is a great predictor for overall SBU rates

Question? Use the Webex chat facility to ask the Host

Performance Relative to Westmere



- Performance numbers plotted against date of first availability of SGI ICE product with that Xeon node type
- Notes:
  - *Clovertown SBU point derived from a different SBU app suite, and is a high-bin part rather than the mid-bin used in Pleiades
  - **Sandy Bridge numbers are preliminary and subject to revision

Question? Use the Webex chat facility to ask the Host

Which is Best Platform for an Application?

- Experiment! Run on each Xeon type:
  - Use as many cores per node as you can
    http://www.nas.nasa.gov/hecc/support/kb/Resources-Request-Examples_188.html
  - If there is enough memory, you can even try Hyper-Threading
    http://www.nas.nasa.gov/hecc/support/kb/Nehalem-EP-Processors_79.html
- Do the math—choose platform that minimizes SBUs:

  (*Nodes used* * *runtime* * *charging factor*)
- For example, OVERFLOW (480 ranks) as used in SBU determination:

	Harpertown	Nehalem	Westmere	Sandy Bridge	
# ranks/node	8	8	12	16	
Number of nodes	60	60	40	30	
Runtime (hr)	0.81	0.38	0.5	0.32	0.35
SBU rate	0.4	0.8	1.0	1.82	1.65
Total SBU cost	19.54	18.19	19.84	17.50	17.23

Question? Use the Webex chat facility to ask the Host